ISSN: 2476-6909; Modares. Mech. Eng. *2025;25(09):587-594.*

Modares Mechanical Engineering

DOI: 10.48311/mme.2025.27590
journal homepage: mme.modares.ac.ir

Surface Roughness and Microhardness Prediction: A Machine Learning Approach for Electrochemical Grinding

Amir Rasti[1*], Amir Hossein Rabiee[2], Ali Zeinolabedin-Beygi[1], Mohammad Yazdani[1]

[1] Advanced Technology of Machine Tools Laboratory (ATMT), Faculty of Mechanical Engineering, Tarbiat Modares University, Tehran, Iran.
[2] Department of Mechanical Engineering, Arak University of Technology, Arak, Iran.

ARTICLE INFO

Article Type
Original Research

Article History
Received: July 20, 2025
Accepted: October 04, 2025
ePublished: November 01, 2025

ABSTRACT

This study focuses on the development of predictive machine learning models to estimate key surface integrity parameters in the electrochemical grinding (ECG) of AISI 304 stainless steel. Experimental data were collected from a series of 20 controlled tests based on a response surface methodology (RSM) design, varying three primary process parameters: voltage, electrolyte concentration, and grinding wheel speed. Using this dataset, Gaussian Process Regression (GPR) models were constructed for four output variables: current density, surface roughness in X- and Y-directions (Ra_x and Ra_y), and surface microhardness (Vickers). Model performance was evaluated using R^2 scores, residual analysis, and error distributions across both training and test datasets. The results demonstrate that surface roughness parameters, particularly Ra_y (R^2_test = 0.970) and Ra_x (R^2_test = 0.932), were predicted with the highest accuracy and consistency. Current density also exhibited strong performance (R^2_test = 0.954), though with minor deviations at extreme values. Surface microhardness, in contrast, posed greater modeling challenges, achieving the lowest test R^2 (0.843) and showing systematic underprediction. Residual and error analyses confirmed these trends, with minimal bias and variance for Ra_x and Ra_y, and broader, asymmetric error profiles for hardness. The least roughness was observed under an electrolyte concentration of 140 g/L, an applied voltage of 20 V, and a grinding wheel rotational speed of 2000 rpm. Overall, the GPR models proved effective for capturing ECG process behavior and offer potential for process optimization in precision manufacturing.

Keywords: Electrochemical Grinding, Surface Integrity, Machine Learning, Gaussian Process Regression.

How to cite this article
Rasti A, Rabiee A.H, Zeinolabedin-Beygi A, Yazdani M, Surface Roughness and Microhardness Prediction: A Machine Learning Approach for Electrochemical Grinding. Modares Mechanical Engineering; *2025;25(09):587-594*.

**Corresponding author's email: a.rasti@modares.ac.ir*
**Corresponding ORCID ID: 0000-0002-0138-6764*

1- Introduction

AISI 304 stainless steel is extensively employed across numerous sectors due to its notable resistance to corrosion, satisfactory mechanical strength, and good formability [1]-3]. A critical area of its application lies in the production of medical devices such as surgical instruments, implants, and particularly hypodermic needles [4], 5]. Nevertheless, traditional machining processes for this alloy frequently lead to suboptimal surface finishes, accelerated tool degradation, and reduced performance of medical components. Hypodermic needles, for instance, necessitate exceptional dimensional accuracy and immaculate surface quality to minimize tissue trauma, enhance patient comfort, and lower the risk of infection. Achieving a finely sharpened needle tip with a polished surface demands a manufacturing approach capable of delivering high precision and excellent surface characteristics. Conventional grinding techniques often face limitations in fulfilling these stringent criteria due to problems such as excessive thermal loads, residual stresses, and burr formation.

Electrochemical Grinding (ECG) has proven to be a highly efficient method for addressing the limitations of conventional machining. This process integrates electrochemical dissolution with conventional abrasive grinding, allowing precise material removal while significantly reducing thermal and mechanical stresses. In ECG, over 80% of the material at the workpiece interface is removed through electrochemical action, while the abrasive wheel simultaneously eliminates the remaining material and any modified surface layer, thereby maintaining a steady and smooth material removal rate (MRR) [6]. The inherently non-contact electrochemical mechanism greatly reduces tool wear and effectively prevents burr formation, making ECG particularly advantageous for fabricating medical components where surface integrity is paramount.

Similar to other material removal techniques, ECG relies on a range of vital process parameters that must be carefully adjusted to obtain optimal performance. Among the most influential variables are the type and concentration of the electrolyte, applied voltage, current density, workpiece feed rate, and the specifications of the grinding wheel. The electrolyte—generally a conductive liquid—directly affects the electrochemical reaction kinetics, thereby impacting both the MRR and the resulting surface finish. Voltage and current density regulate the extent of electrochemical dissolution, whereas the feed rate governs the degree of abrasive action and the overall machining productivity. Additionally, the grinding wheel's attributes—such as abrasive material, grain size, and bonding method—play a pivotal role in determining process outcomes [7]. Proper optimization of these factors is critical for minimizing defects like excessive surface roughness, thermal-induced alterations, and premature tool wear, thus ensuring high surface integrity and precise dimensional control of the final parts [8]. This is particularly important when grinding the tips of AISI 304 stainless steel needles for medical use, where stringent quality requirements must be met to guarantee the functionality and safety of the devices. In the subsequent section, existing research related to the surface quality achieved by ECG is discussed.

Huyan et al. [9], for instance, carried out experimental research on the ECG process applied to Inconel 718. Their investigation explored the effects of various operational parameters—including applied voltage, electrical current, electrolyte concentration, and wheel rotational speed—on the MRR. Their findings indicated that, when other variables were kept constant, an increase in current led to a corresponding rise in MRR. Maksoud and Brooks [10] explored the ECG process utilizing grinding wheels embedded with diamond abrasives and bonded with metallic matrices. Their study emphasized the superior performance of ECG compared to traditional grinding methods, highlighting benefits such as higher MRRs and diminished thermal loads. They further reported that the grinding wheels maintained their cutting ability over extended use without frequent dressing, indicating a notable reduction in wheel consumption. Hansaog et al. [11] examined material removal mechanisms during ECG of GH4169 alloy, employing tooling systems with internal channels to supply the electrolyte directly to the machining zone. Their findings revealed that the mode of electrolyte delivery significantly influences the process stability and overall efficiency. Experimental outcomes demonstrated that increased diamond grit size, higher applied voltage, elevated electrolyte pressure, and higher temperatures contributed to improvements in feed rate and MRR. Yehia et al. [12] studied the influence of Al_2O_3 powder additions on the MRR and surface finish during ECG. Their primary objective was to enhance MRR and improve surface integrity under high-voltage conditions. Levinger and Malkin [13] applied ECG to tungsten carbide-cobalt (WC-Co) materials, reporting that the specific mechanical energy required per unit volume removed by ECG was considerably lower than that required for conventional grinding. Li et al. [14] conducted experiments on ultrasonic-assisted ECG (UAECG) for Ti-6Al-4V alloys. To validate the feasibility of UAECG, they developed a custom experimental setup and characterized its fundamental machining performance. They reported that UAECG led to reductions of about 57% and 56% in normal and tangential forces, respectively. They also found that wheel life in ECG and UAECG was primarily affected by chip adhesion and abrasive grain fracture, whereas in conventional grinding, wheel wear was mainly due to grain detachment. Hakima et al. [15] investigated AECG on K110 tool steel, analyzing how Al_2O_3 abrasive particles influenced the MRR under various machining parameters, including different voltages (9 and 11.5 volts), feed rates (6.74 to 27.44 inches/min), sodium chloride concentrations (10–20 wt.%), and depths of cut (0 and 0.1 mm). Sapre et al. [16] explored the role of electrolyte flow in micro-ECG operations. They conducted an extensive investigation into how electrolyte flow affects the formation and removal of the oxide layer created during electrochemical dissolution. Their results showed that the dominant contributor to overall MRR was the electrochemical process, followed by abrasive wheel action and erosion by electrolyte flow. Moreover, they observed that cutting forces tended to rise with higher inlet flow rates of the electrolyte.

In recent years, machine learning techniques have found extensive application in forecasting different parameters within the field of mechanical engineering [17]-19]. Rahi and Dubey [20] explored the complex machinability issues of metal matrix composites (MMCs), which arise due to the differing mechanical and thermal properties of matrix and reinforcement phases. They highlighted that while electrochemical machining (ECM) can be applied to MMCs, it often results in poor surface finish and an unwanted oxide layer. To address these drawbacks, they investigaated electrochemical spark grinding (ECSG) as a hybrid alternative. Using an AI-based hybrid ANN-GA model, they successfully predicted and optimized surface roughness, oxide layer thickness, and micro-crack width, achieving notable improvements in overall surface quality. Cebi et al. [21] investigated robotic electrochemical machining (RECM) for freeform surfaces, showing that MRR increases with voltage, electrolyte conductivity, and feed rate, similar to conventional ECM. They developed three machine learning models with DNN outperforming others in predicting machining outcomes. RECM demonstrated uniform machining and improved surface roughness, highlighting its potential for complex materials.

Despite the critical importance of achieving superior surface integrity in the manufacturing of medical components such as surgical needles, existing research on ECG still faces significant gaps. Most previous studies have focused primarily on analyzing ECG process parameters through conventional experimental methods without extending these findings toward predictive frameworks for process optimization. Furthermore, while ECG has proven to be highly effective for machining hard-to-cut materials with minimal thermal damage and burr formation, little effort has been devoted to developing reliable models that can accurately estimate key surface integrity outcomes under varying process conditions. Although machine learning

approaches have shown considerable promise in other advanced manufacturing contexts, their application in ECG—particularly for predicting multi-output responses like surface roughness in different directions, current density, and microhardness—remains largely unexplored. In addition, prior works often overlook the combined influence of multiple controllable parameters using systematic experimental designs that ensure comprehensive data coverage for robust modeling. To address these limitations, the present study integrates a structured response surface methodology (RSM) with modern data-driven modeling to provide an effective and practical tool for predicting surface integrity in ECG of AISI 304 stainless steel. By generating a reliable experimental dataset and leveraging it for model development, this research aims to contribute to more precise process control and enhanced quality assurance in critical medical manufacturing applications.

2- Material and Methods

The primary objective of this research was to generate reliable experimental data to develop predictive models based on machine learning for estimating key surface integrity parameters during the ECG of AISI 304 stainless steel. For this purpose, a comprehensive laboratory-scale ECG setup was custom-designed to enable precise control over critical process parameters and ensure repeatable machining conditions. The test rig consisted of a spindle drive equipped with a 120 mm diameter cup-shaped Tyrolit grinding wheel, an adjustable electrolyte delivery system, a DC power supply, and a one-dimensional feed control mechanism. The grinding wheel, manufactured with a metallic bonding structure containing Al_2O_3 abrasive grains (grit size 125), performed the dual role of facilitating mechanical material removal while simultaneously conducting electric current to enable the electrochemical reaction. The electrolyte used was a saline solution prepared by mixing distilled water with sodium chloride, delivered at a constant flow rate of 5 L/min using a SUBO WP-3200 pump to maintain continuous flushing of the machining zone, effective heat removal, and stable electrochemical dissolution. Figure 1 shows the experimental setup. The workpieces were rectangular blocks of AISI 304 stainless steel with dimensions of 6 × 6 × 60 mm and an average hardness of 130 ± 3 Hv. Prior to and after each experiment, the samples were ground, ultrasonically cleaned, and weighed with a precision balance to verify consistency in material removal measurements. To study the influence of main processing variables, three parameters—electrolyte concentration, applied voltage, and grinding wheel rotational speed—were selected and systematically varied across five levels, as detailed in Table 1, using a central composite RSM design. A total of 20 experiments were conducted, and the complete experimental matrix along with the recorded output data are presented in Table 2. Surface roughness (Ra) in both X- and Y-directions was measured using a Mahr-Marsurf PS1 roughness meter with a 1.75 mm evaluation length. To minimize local variations, measurements were repeated at three separate points on each workpiece and averaged. Microhardness of the machined surfaces was also determined using a Buehler hardness tester under a load of 0.1 kg for 4 seconds. The resulting dataset, consisting of current density, Ra_x, Ra_y, and surface microhardness, was extracted with the goal of training and validating predictive models for accurate surface integrity prediction in ECG processes.

3- Machine Learning Model

In this study, Gaussian Process Regression (GPR) was employed to model and predict the relationship between key ECG process parameters, namely electrolyte concentration, voltage, and wheel speed, and four surface integrity outputs: current density, surface roughness in X and Y directions (Ra_x and Ra_y), and microhardness. GPR is a non-parametric, Bayesian approach to regression that offers flexibility in capturing nonlinear behaviors and uncertainty estimation, making it highly suitable for small-scale experimental datasets with complex physical interactions.

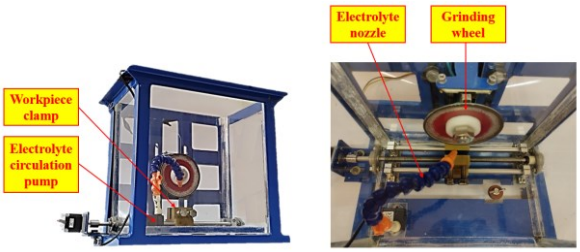

Fig.1 Experimental setup for ECG process

Table 1 Levels of process parameters used in ECG experiments [22]

Process parameters	Levels				
	-2	-1	0	1	2
Grinding wheel rotational speed (rpm)	500	1000	1500	2000	2500
Electrolyte concentration (g/L)	20	60	100	140	180
Applied voltage (V)	5	10	15	20	25

Table 2 Design matrix and measured surface integrity responses [22]

No.	Input			Output			
	Electrolyte concentration (g/L)	Applied voltage (V)	Grinding wheel rotational speed (rpm)	Current density (A/cm^2)	Ra_y (µm)	Ra_x (µm)	H (Hv)
1	20	15	1500	200.3	0.848	0.709	138
2	60	10	1000	310.8	0.693	0.690	134
3	60	10	2000	138.7	0.656	0.696	145
4	60	20	1000	457.8	0.616	0.473	126
5	60	20	2000	374.5	0.609	0.473	134
6	100	5	1500	155.5	0.680	0.658	137
7	100	15	500	549.8	0.588	0.524	129
8	100	15	1500	452.2	0.423	0.384	134
9	100	15	1500	446.1	0.416	0.382	134
10	100	15	1500	450.2	0.412	0.382	133
11	100	15	1500	448.0	0.434	0.383	132
12	100	15	1500	442.3	0.449	0.381	134
13	100	15	1500	448.1	0.439	0.382	132
14	100	15	2500	305.3	0.402	0.535	142
15	100	25	1500	727.6	0.523	0.462	118
16	140	10	1000	448.7	0.473	0.413	134
17	140	10	2000	287.3	0.345	0.419	135
18	140	20	1000	813.5	0.421	0.434	121
19	140	20	2000	705.5	0.306	0.440	127
20	180	15	1500	683.4	0.352	0.392	124

A Gaussian Process (GP) is a collection of random variables, any finite number of which have a joint Gaussian distribution. In regression problems, GPR models the function $f(x)$ mapping input variables $\boldsymbol{x} \in \mathbb{R}^d$ to outputs as a distribution over functions, as shown in Equation 1 [23]:

$$f(x) \sim GP(m(x), k(x, x')) \quad (1)$$

Where $m(x)$ is the mean function, typically assumed to be zero, and $k(x, x')$ is the kernel (or covariance function), which defines the covariance between any two input points.

3-1- Kernel Selection

In this study, a Matérn kernel with $\nu = 1.5$ was chosen due to its balance between flexibility and smoothness. The complete kernel used is shown in Equation 2 [24]:

$$k(\boldsymbol{x}, \boldsymbol{x}') = \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{l}\right) exp\left(-\frac{\sqrt{3}r}{l}\right) \quad (2)$$

Where $r = \| x - x' \|$ is the Euclidean distance between input vectors, l is the length scale, and σ_f^2 is the signal variance. The Matérn kernel with $\nu = 1.5$allows modeling functions that are once differentiable, suitable for physical processes where abrupt changes (e.g., due to threshold effects in current density or corrosion) may occur.

3-2- Model Training and Optimization

Each output variable was independently modeled using a separate GPR. The dataset was initially partitioned into a training set comprising 70% of the data and a test set containing the remaining 30%. The kernel hyperparameters were optimized internally via maximum likelihood estimation, which ensures robustness against local minima.

The predicted outputs $\hat{y}$ were computed from the posterior distribution, as shown in Equation 3 [25]:

$$\hat{f}(x_*) = k_*^T (K + \sigma_n^2 I)^{-1} y \quad (3)$$

where, K is the covariance matrix of training inputs, k_* is the covariance vector between test and training points, σ_n^2 is the noise variance (implicitly estimated).

3-3- Model Training and Optimization

Length scale l controls the smoothness of the function. Smaller values allow for rapid changes in the function output but can lead to overfitting, especially in noisy data such as hardness measurements. Signal variance σ_f^2 governs the amplitude of function variations. A higher value indicates greater model flexibility but must be regularized to avoid capturing noise. Number of optimizer restarts increasing Number of optimizers helps escape poor local minima in marginal likelihood optimization and was set to 10 in all models to ensure reliable convergence. The hyperparameters of the Gaussian Process Regression models (length scale, signal variance, and noise variance) were optimized through maximum likelihood estimation (MLE), ensuring stable and reliable parameter tuning.

3-4- Model Evaluation

For each output, the trained model's performance was evaluated using standard regression metrics: R^2 (coefficient of determination), RMSE (Root Mean Squared Error), and MAE (Mean Absolute Error), as shown in Equations 4-6 [26]:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_i (y_i - \hat{y}_i)^2} \quad (5)$$

$$\text{MAE} = \frac{1}{n} \sum_i | y_i - \hat{y}_i | \quad (6)$$

Where $\hat{y}_i$ is the predicted value and y_i is the true target. By leveraging GPR with Matérn kernel and careful hyperparameter tuning, the study achieved high-accuracy predictive models that effectively capture nonlinear relationships in ECG surface integrity data. The framework demonstrated here is particularly useful for small experimental datasets where uncertainty quantification and nonlinearity are critical.

4- Results and Discussion

Based on the previous methodology, the GPR models were configured with a composite kernel consisting of a Constant Kernel multiplied by a Matérn kernel. The constant kernel had an initial variance parameter of 1.0, while the Matérn kernel was specified with a length scale of 1.0 and smoothness parameter $\nu = 1.5$, which allows for once-differentiable functions—suitable for capturing moderate nonlinearity in ECG data. Additionally, the hyperparameter number of optimizers was set to 10 to ensure stable convergence during marginal likelihood optimization and to avoid poor local minima.

Figure 2 shows the predicted versus actual values for all four output variables (Current Density, Ra_y, Ra_x, and Hardness) based on the GPR models. Each subplot corresponds to one output, where the blue circles represent the training samples and the red squares denote the test samples. The dashed diagonal line indicates perfect prediction, where predicted values match the actual ones exactly. For Ra_y and Ra_x, the predicted points lie very close to the diagonal, both for training and test sets, demonstrating that the GPR model has learned the underlying surface roughness patterns with high accuracy. In particular, the predictions for Ra_y align almost perfectly with the true values, indicating strong generalization, which is consistent with its highest R^2 on the test set (0.970). Similarly, Ra_x shows tight prediction accuracy with minimal deviation, confirming its R^2_test of 0.932.

In contrast, the Hardness subplot reveals more dispersion, especially in the test data. Several red squares deviate from the diagonal line, particularly at the lower and mid hardness ranges, indicating a slight drop in model performance. This corresponds to its lower R^2_test of 0.843, which, although still acceptable, suggests that surface hardness is more complex to model with GPR, likely due to nonlinearities or unmodeled interactions. The Current Density plot exhibits overall good agreement with the true values, especially in the mid-range, but test predictions tend to deviate slightly more at the extremes. The model maintains a high R^2_test of 0.954, but the slight under- and over-predictions near the boundaries reduce its comparative reliability in extrapolating extreme values. Among all outputs, Ra_y demonstrates the best model performance, followed closely by Ra_x and Current Density. Hardness is the least accurately predicted parameter, as seen in the larger spread of points in its test data. This figure provides a clear visual comparison of model fidelity across different outputs, highlighting areas of both strong and moderate predictive accuracy.

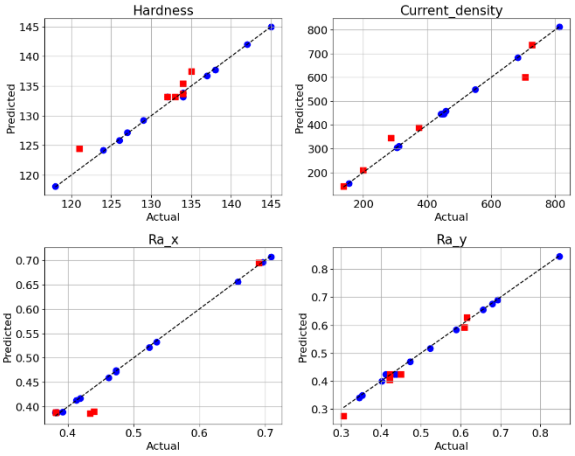

Fig.2 Predicted vs. actual values for four output parameters: Current Density, Surface Roughness in X-direction (Ra_x), Surface Roughness in Y-direction (Ra_y), and Surface Microhardness. The dashed diagonal line represents the ideal case where predicted equals actual. Blue circles (●) indicate training data, and red squares (■) represent test data

Figure 3 illustrates the distribution of prediction errors for each output variable using histograms. These plots provide insights into the magnitude and symmetry of the errors generated by the GPR models across the entire dataset.

The error distribution for Ra_x is sharply concentrated around zero, indicating extremely low deviation between predicted and actual values. This narrow and symmetric distribution suggests high accuracy and stability of the model in estimating surface roughness along the X-direction. Ra_y displays a similarly narrow spread, although with slightly more variation, confirming its strong predictive behavior with minimal bias.

Current Density shows a broader error distribution, with most errors clustered near zero but a few outliers on both sides. This pattern indicates that while the model performs well in most cases, it occasionally underestimates or overestimates the current density, especially in regions where the experimental behavior is more nonlinear or extreme.

The Hardness histogram displays the widest spread among all outputs, with several negative error values reaching below -3. This suggests that the model tends to underpredict hardness in certain conditions, especially in test samples with lower or variable hardness. The asymmetry in this distribution reflects the model's limited ability to generalize well across the full hardness range, as also observed in the previous figure.

By comparing these histograms, it becomes clear that the models predicting Ra_x and Ra_y exhibit the most consistent and precise performance, while the model for Hardness, although generally accurate, has a higher error variability. Current Density falls in between, with moderate dispersion and a few significant outliers. These distributions support the earlier conclusion that surface roughness parameters are more reliably predicted by the GPR model compared to hardness and current density.

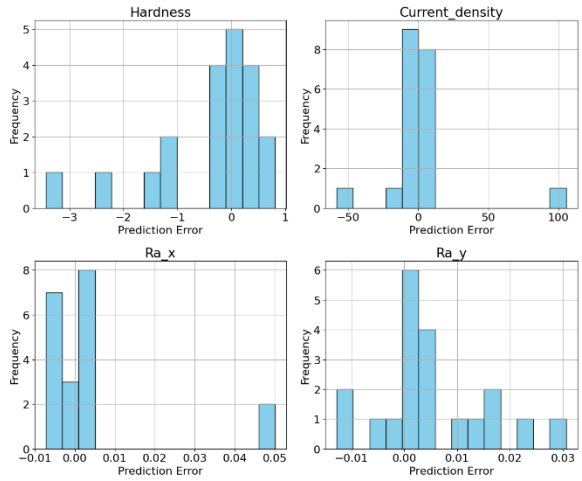

Fig.3 Histogram plots of prediction errors for each output parameter: Surface Microhardness, Current Density, Surface Roughness in X-direction (Ra_x), and Y-direction (Ra_y). The horizontal axis represents the prediction error (Actual − Predicted), and the vertical axis shows the frequency of occurrences

Figure 4 displays the residual plots for the four predicted output variables against the predicted values from the GPR models. Residuals, defined as the difference between actual and predicted values, are plotted on the vertical axis, while the corresponding predicted values are on the horizontal axis. These plots help assess whether the model errors are randomly distributed, which is a key assumption in well-behaved regression models.

For Ra_y and Ra_x, the residuals are tightly scattered around zero and exhibit no visible trend across the range of predicted values. This indicates that the GPR models for surface roughness in both directions perform consistently and do not systematically over- or under-predict for specific value ranges. The narrow spread of residuals for Ra_x, in particular, confirms its robustness and lack of bias throughout the prediction range.

In the case of Current Density, the residuals show a slightly wider band, especially at the higher end of predicted values. There is a slight tendency for underestimation at high current densities, as evidenced by a mild downward curvature. However, the residuals remain centered around zero, suggesting that the model still retains acceptable performance, albeit with slightly increasing variance in extreme values.

The residual plot for Hardness reveals a more scattered pattern, with residuals showing asymmetry and greater variance at both ends of the prediction range. In some areas, especially where predicted hardness is high, the residuals become predominantly negative, indicating a tendency to underpredict. This behavior aligns with the earlier findings that the model has lower generalization performance for surface hardness.

Overall, these residual plots reinforce the conclusion that Ra_x and Ra_y are the most accurately and stably predicted outputs, followed by Current Density. Hardness remains the most difficult parameter for the model to predict reliably across the entire value range.

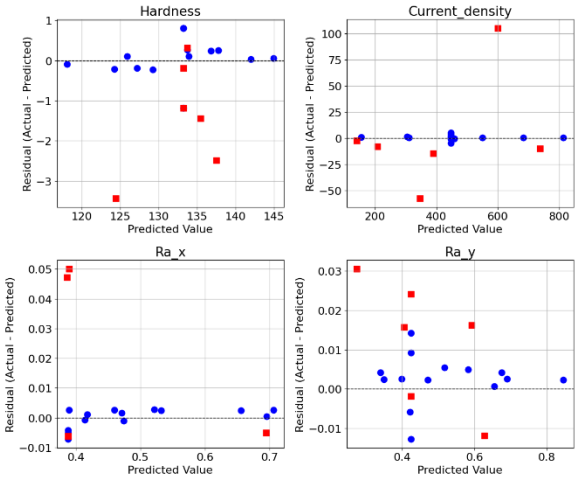

Fig.4 Residual plots of GPR predictions for four output parameters: Current Density, Surface Roughness in Y-direction (Ra_y), Surface Roughness in X-direction (Ra_x), and Surface Microhardness. The vertical axis shows the residuals (Actual − Predicted) and the horizontal axis the predicted values

Residual analysis revealed that the very narrow clustering of Ra_x and Ra_y residuals, while demonstrating high accuracy, may also indicate mild overfitting given the limited dataset size. Current density residuals showed slight curvature at high values, suggesting reduced extrapolation capability due to partial overfitting in mid-range regions. In contrast, hardness residuals reflected greater variability and systematic underprediction, consistent with underfitting rather than overfitting.

Figure 5 displays the actual and predicted values for all four output variables (Current Density, Hardness, Ra_y, and Ra_x) across individual sample indices, separated into training and test sets. The light blue and red lines represent the true values for training and test data, respectively, while the blue circles and red squares denote the corresponding predicted values from the GPR models. This figure provides a temporal or sequential perspective of model accuracy across the full dataset. In the case of Current Density, the predicted values in the training set closely follow the actual values, showing minimal deviation. For the test set, although the model captures the overall trend, it slightly overestimates or underestimates a few points, particularly at sample indices 16 and 18, where the predicted values deviate from the actual spikes. This behavior indicates a strong fit overall, but with reduced performance on high-gradient regions.

The plot for Hardness reveals more significant prediction error in the test samples. Although the model follows the general trend of the test data, the predicted values exhibit a consistent downward shift, particularly at indices 15 through 17. This suggests a systematic underprediction of hardness values in that range, confirming that hardness remains the most difficult target for the model to generalize across unseen data. For Ra_y, the model shows excellent agreement between predicted and actual values throughout both training and test sets. The predicted curve tightly follows the actual one with minimal deviation, reflecting the high R^2_test and confirming the model's

capability in capturing the underlying structure of surface roughness in the Y-direction. Ra_x also demonstrates strong performance in the training set, with predicted values tracking closely with the actual measurements. However, in the test set, the predicted values become almost constant, failing to capture the variation in actual surface roughness at sample indices 15 through 20. This flattening effect indicates that while the model fits well in the training phase, it may be oversimplifying in the test region, possibly due to limited variance in the input features for those samples. Altogether, this figure reinforces previous findings that Ra_y is the most reliably predicted output, followed by Ra_x and Current Density. Hardness continues to show the highest prediction errors in the test set, especially at specific indices, reflecting the model's challenge in extrapolating to regions with more complex behavior.

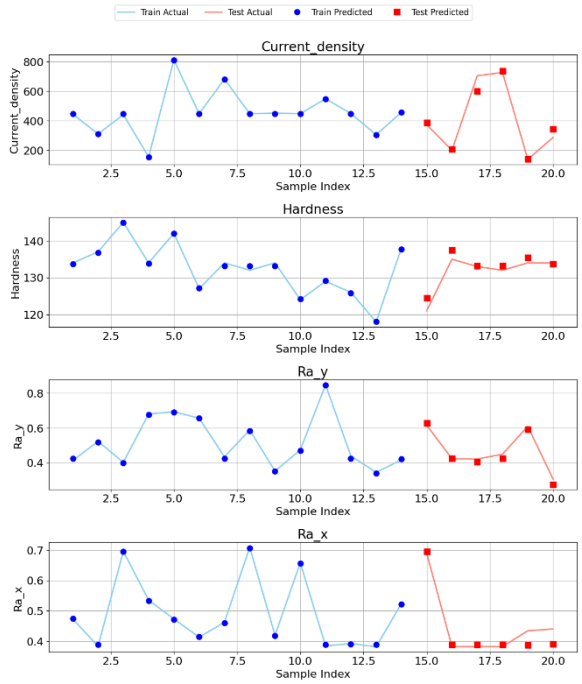

Fig.5 Line plots of actual and predicted values across all samples for each output parameter: Current Density, Surface Microhardness, Ra_y, and Ra_x. Light blue lines show actual values in the training set, and red lines show actual values in the test set. Blue circles (●) and red squares (■) represent model predictions for training and test data, respectively

Figure 6 presents box plots of the prediction errors for each output parameter separately for the training and test datasets. The vertical axis represents the prediction error, defined as the difference between the actual and predicted values. These box plots provide a visual summary of the error distribution, central tendency, and the presence of outliers, allowing for a direct comparison of model performance between training and test phases. For Hardness, the training error distribution is narrow and centered around zero, indicating a well-fitted model. However, the test error box is noticeably shifted downward and significantly wider, with several test samples underpredicted by up to 3 units. This reveals a clear drop in generalization ability for Hardness, consistent with its relatively low R^2_test, and highlights that the model has difficulty extrapolating hardness values beyond the training domain. In the case of Current Density, the training error is tightly concentrated with minimal spread, while the test error shows a wider distribution and a negative bias, suggesting a slight tendency to underpredict current density on unseen data. The presence of outliers, particularly one with a large positive error, indicates isolated mispredictions, potentially linked to the extremes of the parameter space.

The Ra_x box plots show excellent model stability during training, with almost no variance. Although the test error spread is broader, it remains relatively small and symmetric. Importantly, the median error is very close to zero, confirming that even with increased variability in the test set, the model maintains its predictive accuracy and does not exhibit systemic bias. Similarly, Ra_y demonstrates a narrow and centered training error, and although the test error box is slightly more spread out, it still remains close to zero with minimal outliers. This reflects a strong balance between training fit and test generalization, further validating Ra_y as the most consistently predicted output among all parameters. Overall, these box plots clearly show that surface roughness metrics (Ra_x and Ra_y) are modeled with high consistency and robustness, while Current Density retains acceptable performance with minor deviations. In contrast, Hardness emerges as the least stable prediction target, exhibiting both higher variance and directional bias in the test set.

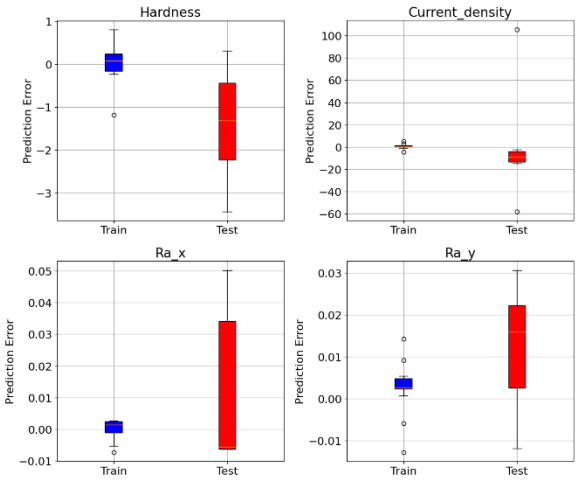

Fig.6 Box plots of prediction errors for each output variable—Surface Microhardness, Current Density, Surface Roughness in X-direction (Ra_x), and Y-direction (Ra_y)—separated for training and test datasets. The vertical axis shows the prediction error (Actual − Predicted). Blue boxes represent training data and red boxes test data

Compared to Ra_x/Ra_y and current density, microhardness exhibited a lower test R^2 and a mild underprediction. Physically, hardness is affected by additional, unmeasured factors (e.g., electrolyte temperature/flow non-uniformity, passive-film dynamics, subsurface strain hardening, residual stress/microstructural changes), which weakens the mapping from the three controlled inputs. Statistically, the narrower hardness range and higher local measurement variability (micro-indentation) reduce the signal-to-noise ratio and yield asymmetric, heteroscedastic errors. With 20 experiments and single-output GPR, potential cross-output structure remains unexploited; furthermore, a moderately smooth Matérn kernel may underfit sharp electrochemical transitions. These factors jointly explain the lower predictive fidelity for hardness.

Table 3 summarizes the predictive performance of the machine learning models developed for each output parameter derived from the electrochemical grinding experiments. This numerical summary complements the visual analysis provided in Figures 2 through 6 and reinforces the key conclusions drawn regarding model accuracy and generalization capability. Among the four predicted outputs, Ra_y achieved the highest test R^2 value (0.970), demonstrating excellent model performance and strong generalization to unseen data. This observation is consistent with earlier residual and error distribution plots, where Ra_y showed tightly clustered and unbiased predictions. Ra_x also exhibited high accuracy, with a test R^2 of 0.932, indicating that the model captured the surface roughness trends in the X-direction effectively, although some flattening behavior was observed in test predictions for certain indices. The Current Density model achieved a solid R^2_test of 0.954, with generally consistent predictions and only

minor deviations, primarily in high-value regions as identified in Figure 5.

Hardness was the most challenging parameter to predict, with the lowest R²_test (0.843). This relatively lower value reflects the broader error distribution and systematic underestimation observed in both box plots and line charts, suggesting that surface microhardness may involve additional nonlinear interactions or hidden variables not fully captured by the model.

Overall, this table quantitatively confirms that surface roughness parameters, particularly Ra_y, are predicted with the highest reliability, while hardness prediction requires further refinement. The performance metrics presented here validate the effectiveness of GPR models for modeling surface integrity in electrochemical grinding, especially when dealing with multidimensional, small-scale experimental datasets. The trained GPR models can serve as surrogates for ECG process optimization. Applied voltage, electrolyte concentration, and grinding-wheel speed should be adjusted to minimize Ra_y and Ra_x, while maintaining hardness near the experimental baseline and keeping current density within stable mid-range values.

Table 3 Performance metrics of machine learning models used to predict each experimental output parameter. R²_train and R²_test represent the coefficient of determination for training and test sets, respectively. RMSE_test refers to Root Mean Square Error, and MAE_test is Mean Absolute Error

Output	Model	R²_train	R²_test	RMSE_test	MAE_test
Current_density	GPR	0.999794	0.953830	49.774947	33.081041
Ra_y	GPR	0.997970	0.969985	0.019055	0.016749
Hardness	GPR	0.995421	0.842742	1.898375	1.508942
Ra_x	GPR	0.999192	0.932010	0.028526	0.020176

5- Conclusion

This study aimed to develop predictive machine learning models for estimating key surface integrity parameters resulting from ECG of AISI 304 stainless steel. Using experimental data obtained from a controlled DOE framework, GPR models were trained and evaluated to predict current density, surface roughness (Ra_x, Ra_y), and microhardness. The following key findings summarize the predictive performance and insights derived from the models:

- Ra_y was the most accurately predicted output, achieving the highest generalization capability with an R²_test of 0.970. The model consistently tracked actual values across both training and test samples with minimal error and no visible bias in residuals.
- Ra_x also exhibited strong prediction accuracy with an R²_test of 0.932. Although slight flattening occurred in the test predictions, the overall performance remained stable and robust, especially in the training phase.
- Current Density predictions were generally reliable, with a high R²_test of 0.954. However, the model showed a slight tendency to under- or over-estimate extreme values, likely due to nonlinearities not fully captured in the input space.
- Hardness proved to be the most difficult parameter to predict, with the lowest R²_test value (0.843). The prediction errors showed a clear negative bias in the test data and higher variance, suggesting potential hidden or nonlinear factors influencing surface microhardness during ECG.
- Error histograms and box plots confirmed that surface roughness parameters were predicted with the highest precision and lowest variability, while hardness predictions were more dispersed and prone to underestimation.

While GPR is well suited for small-scale, low-dimensional experimental datasets, its cubic complexity and data requirements limit scalability to higher-dimensional problems. For more complex ECG modeling with additional process variables, tree-based ensemble methods (Random Forests, Gradient Boosting) or, with larger datasets, deep neural networks could provide more efficient and scalable alternatives. Multi-output approaches may further improve prediction by exploiting correlations between surface integrity responses.

Ethics Approval

The scientific content of this article is the result of the authors' research and has not been published in any Iranian or international journal.

Conflict of Interest

The authors declare that they have no conflicts of interest to this work.

References

[1] M. K. Firouzjaei, H. M. Naeini, M. M. Kasaei, M. J. Mirnia, and L. F. da Silva, "A microscale constitutive model for thin stainless steel sheets considering size effect," Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications, vol. 237, no. 10, pp. 2104-2114, 2023. doi:10.1177/14644207231169456

[2] M. Karimi Firouzjaei, H. Moslemi Naeini, M. M. Kasaei, and M. J. Mirnia, "A constitutive model for stainless steel 304 sheet considering size effect in micro-scale," Modares Mechanical Engineering, vol. 22, no. 8, pp. 519-528, 2022. doi:10.52547/mme.22.8.519

[3] M. K. Firouzjaei, H. M. Naeini, M. M. Kasaei, M. J. Mirnia, and L. F. da Silva, "Microscale modeling of the ductile fracture behavior of thin stainless steel sheets," Thin-Walled Structures, vol. 196, p. 111457, 2024. doi:10.1016/j.tws.2023.111457

[4] R. Narayan, "Medical application of stainless steels," ASM handbook, vol. 23, pp. 199-210, 2012. doi:10.31399/asm.hb.v23.a0005673

[5] D. E. Wert and R. P. DiSabella, "Strong, corrosion-resistant stainless steel: this strong, tough, corrosion-resistant stainless steel improves performance of equipment in aerospace, medical instruments, oil drilling, firearms, and marine applications," Advanced materials & processes, vol. 164, no. 8, pp. 34-37, 2006.

[6] R. E. Ilhan, Analysis, optimization, and mechanistic modelling of electrochemical surface grinding (ECG) process. Lehigh University, 1991.

[7] D. S. Patel, V. K. Jain, and J. Ramkumar, "Electrochemical grinding," in Nanofinishing science and technology: CRC Press, 2016, pp. 341-372.

[8] P. Thanki, "Electrochemical grinding process, current state and future direction: A systematic literature review," Int. J. Appl. Eng. Res., vol. 9, no. 6, pp. 637-644, 2014.

[9] B. Bhuyan, C. Garg, and L. Gupta, "Design and development of tabletop electrochemical grinding setup," Materials Today: Proceedings, vol. 21, pp. 1479-1482, 2020. doi:10.1016/j.matpr.2019.11.058

[10] T. Maksoud and A. Brooks, "Electrochemical grinding of ceramic form tooling," Journal of materials processing technology, vol. 55, no. 2, pp. 70-75, 1995. doi:10.1016/0924-0136(95)01787-9

[11] L. Hansong, F. Shuxing, Q. Zhang, N. Shen, and Q. Ningsong, "Simulation and experimental investigation of inner-jet electrochemical grinding of GH4169 alloy," Chinese Journal of Aeronautics, vol. 31, no. 3, pp. 608-616, 2018. doi:10.1016/j.cja.2017.08.014

[12] H. M. Yehia, M. Hakim, and A. El-Assal, "Effect of the Al2O3 powder addition on the metal removal rate and the surface roughness of the electrochemical grinding machining," Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, vol. 234, no. 12, pp. 1538-1548, 2020. doi:10.1177/0954405420921721

[13] R. Levinger and S. Malkin, "Electrochemical grinding of WC-Co cemented carbides," 1979.

[14] S. Li, Y. Wu, M. Nomura, and T. Fujii, "Fundamental machining characteristics of ultrasonic-assisted electrochemical grinding of Ti–6Al–4V," Journal of

Manufacturing Science and Engineering, vol. 140, no. 7, p. 071009, 2018. doi:10.1115/1.4039855

[15] M. Hakima, M. Yehiab, A. El-Assal, and K. Abdelwahed, "Hybrid Abrasive Electrochemical Grinding Machining of alloy steel K110," Engineering Research Journal, vol. 163, pp. 294-308, 2019. doi:10.21608/erj.2019.122540

[16] P. Sapre, A. Mall, and S. S. Joshi, "Analysis of electrolytic flow effects in micro-electrochemical grinding," Journal of manufacturing science and engineering, vol. 135, no. 1, p. 011012, 2013. doi:10.1115/1.4023266

[17] A. Zeinolabedin-Beygi, H. M. Naeini, H. Talebi-Ghadikolaee, A. H. Rabiee, and S. Hajiahmadi, "Predictive modeling of spring-back in pre-punched sheet roll forming using machine learning," The Journal of Strain Analysis for Engineering Design, vol. 59, no. 7, pp. 463-474, 2024. doi:10.1177/03093247241263685

[18] H. Ziaiefar, M. Amiryan, M. Ghodsi, F. Honarvar, and Y. Hojjat, "Ultrasonic damage classification in pipes and plates using wavelet transform and SVM," Modares Mechanical Engineering, vol. 15, no. 5, pp. 41-48, 2015.

[19] M. Safari, A. H. Rabiee, and J. Joudaki, "Developing a support vector regression (SVR) model for prediction of main and lateral bending angles in laser tube bending process," Materials, vol. 16, no. 8, p. 3251, 2023. doi:10.3390/ma16083251

[20] D. K. Rahi and A. K. Dubey, "Modeling and optimization of surface quality characteristics in electrochemical surface grinding of metal matrix composite," Journal of Materials Engineering and Performance, vol. 33, no. 9, pp. 4345-4358, 2024. doi:10.1007/s11665-023-08267-9

[21] A. Cebi, H. Demirtas, and A. Kaleli, "Implementation of robotic electrochemical machining in freeform surface machining with material removal rate prediction using different machine learning algorithms," Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, vol. 238, no. 9, pp. 3835-3849, 2024. doi:10.1177/09544062231208302

[22] M. Yazdani and A. Rasti, "Assessment on surface integrity in electrochemical grinding of AISI 304," Heliyon, vol. 11, no. 1, 2025. doi:10.1016/j.heliyon.2024.e41435

[23] C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning (no. 3). MIT press Cambridge, MA, 2006.

[24] M. L. Stein, Interpolation of spatial data: some theory for kriging. Springer Science & Business Media, 1999.

[25] C. Williams and C. Rasmussen, "Gaussian processes for regression," Advances in neural information processing systems, vol. 8, 1995.

[26] H. Talebi-Ghadikolaee, H. Moslemi Naeini, A. H. Rabiee, A. Zeinolabedin Beygi, and S. Alexandrov, "Experimental-numerical analysis of ductile damage modeling of aluminum alloy using a hybrid approach: ductile fracture criteria and adaptive neural-fuzzy system (ANFIS)," International Journal of Modelling and Simulation, vol. 43, no. 5, pp. 736-751, 2023. doi: 10.1080/02286203.2022.2121675